

(12) UK Patent Application (19) GB (11) 2 320 396 (13) A

(43) Date of A Publication 17.06.1998

(21) Application No 9725483.3

(22) Date of Filing 01.12.1997

(30) Priority Data

(31) 9624916

(32) 29.11.1996

(33) GB

(71) Applicant(s)

Northern Telecom Limited

(Incorporated in Canada - Quebec)

World Trade Center of Montreal,
380 St Antoine Street West, 8th Floor, Montreal,
Quebec H2Y 3Y4, Canada

(72) Inventor(s)

Roy Harold Mauger
Simon Daniel Brueckheimer
Brian Michael Unitt

(74) Agent and/or Address for Service

J P W Ryan
Nortel Patents, London Road, HARLOW, Essex,
CM17 9NA, United Kingdom

(51) INT CL⁶

H04L 12/56

(52) UK CL (Edition P)

H4K KTK

(56) Documents Cited

GB 2254529 A

EP 0570630 A1

EP 0552385 A1

(58) Field of Search

UK CL (Edition P) H4K KTK

INT CL⁶ H04L

ONLINE: WPI; INSPEC

(54) Data network router

(57) In a broad band network comprising a plurality of edge switches interconnected by a plurality of tandem switches, IP packet data is routed by determining from each packet whether that packet forms part of a flow of packets. Where a packet is determined to be part of a flow, it is segmented into minicells for transmission over a permanent virtual channel established across the network for that flow. Where a packet is not determined to be part of a flow, a destination edge switch is identified for that packet which is then transported to that edge switch via a default permanent virtual channel. Permanent virtual channels for new flows are established at the network periphery by defining a succession of nodes and virtual channels (VCs) to be used to reach the destination of that flow.

Figure 5

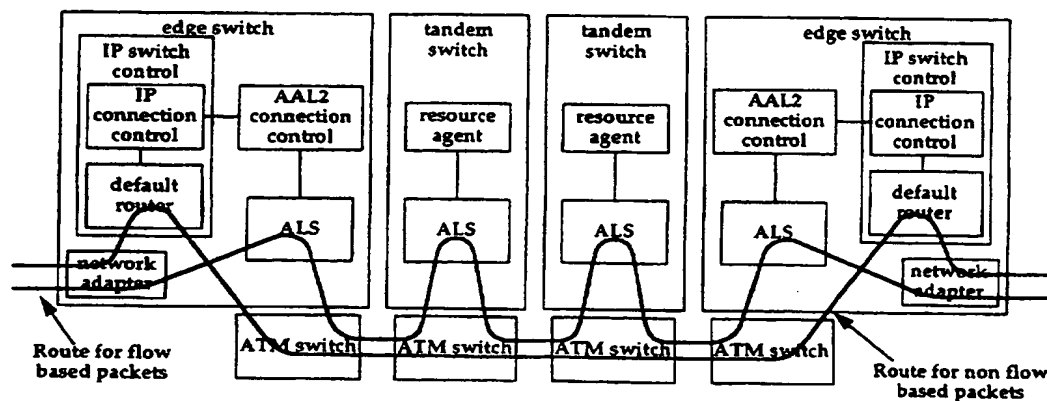
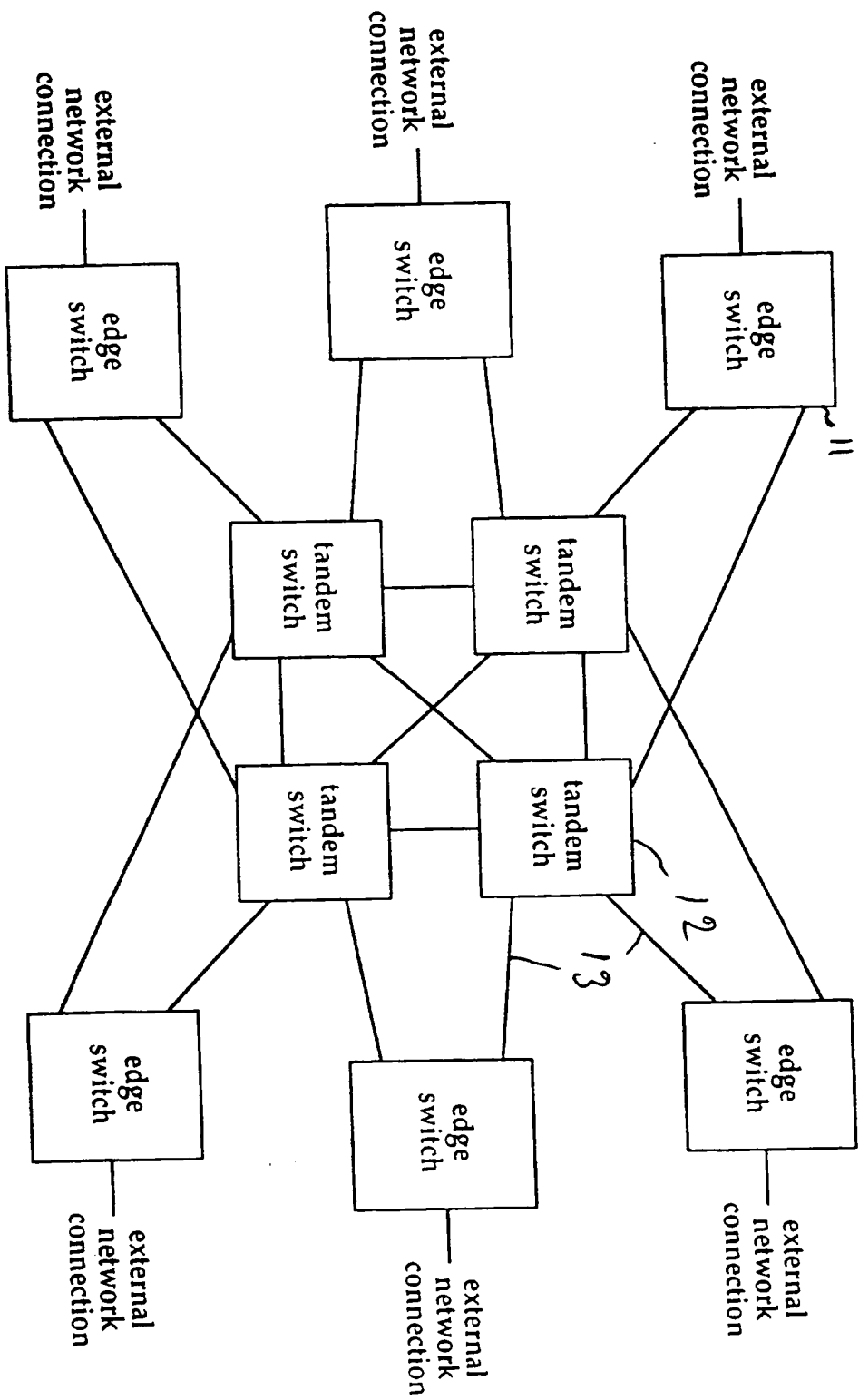


Figure 1

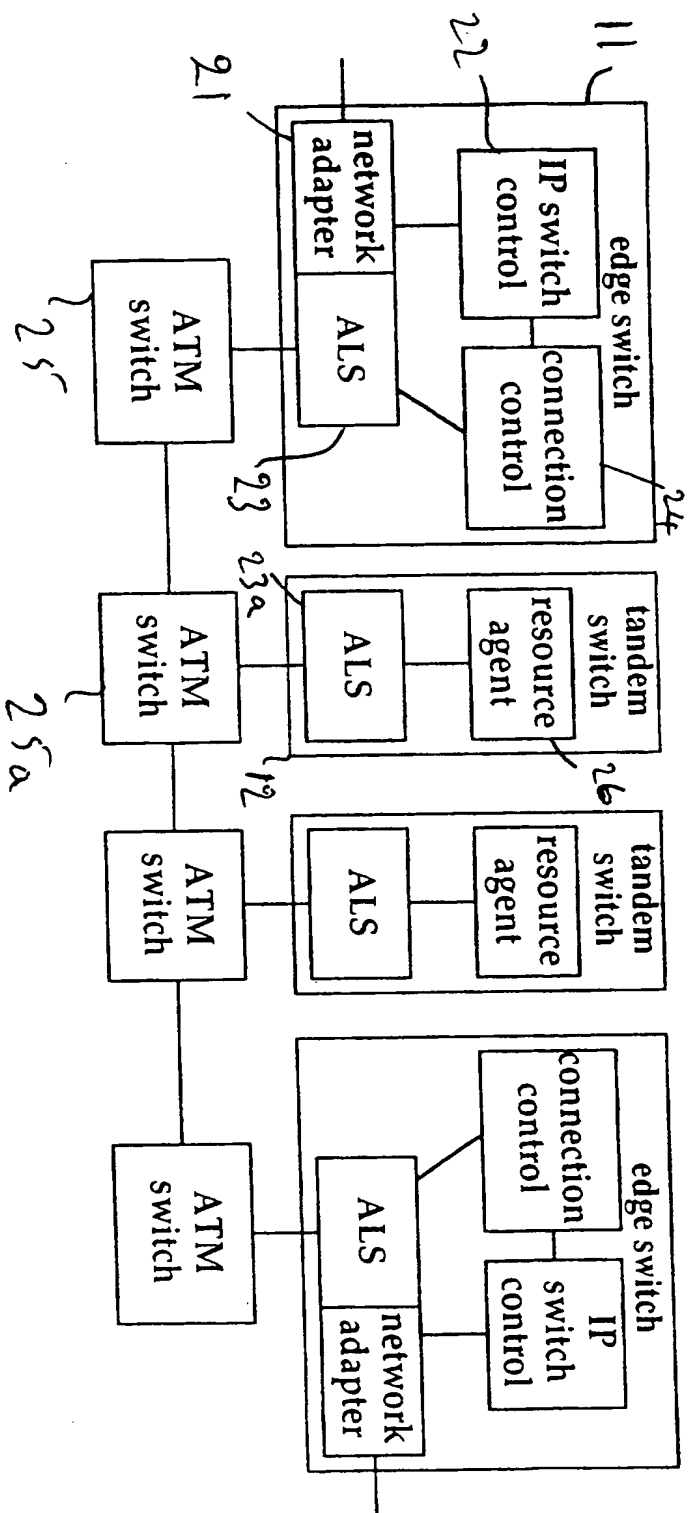


11/6

Figure 2

Tn1, VCm1	Tn2, VCm2	En3, VCm3
-----------	-----------	-----------

Connection Specification "worm"



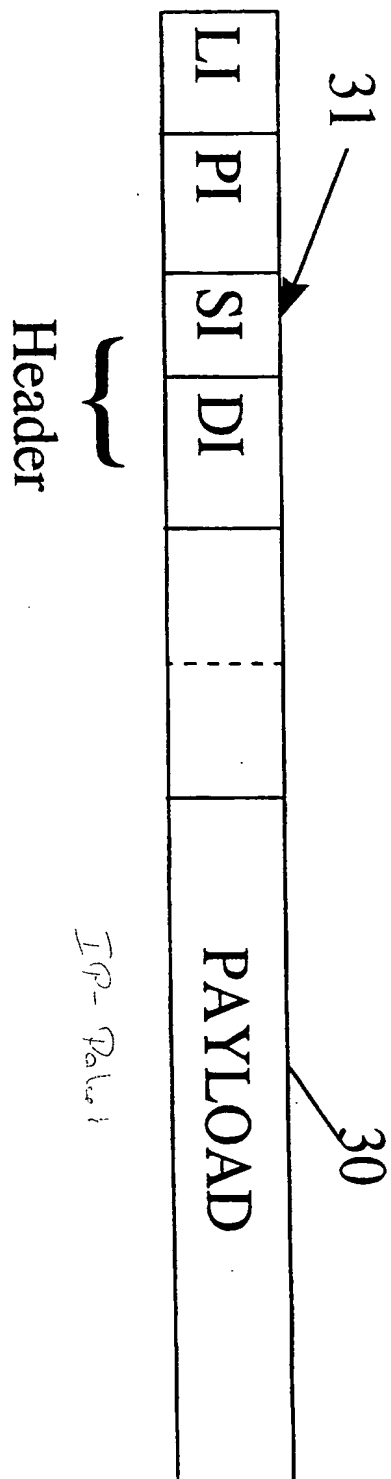
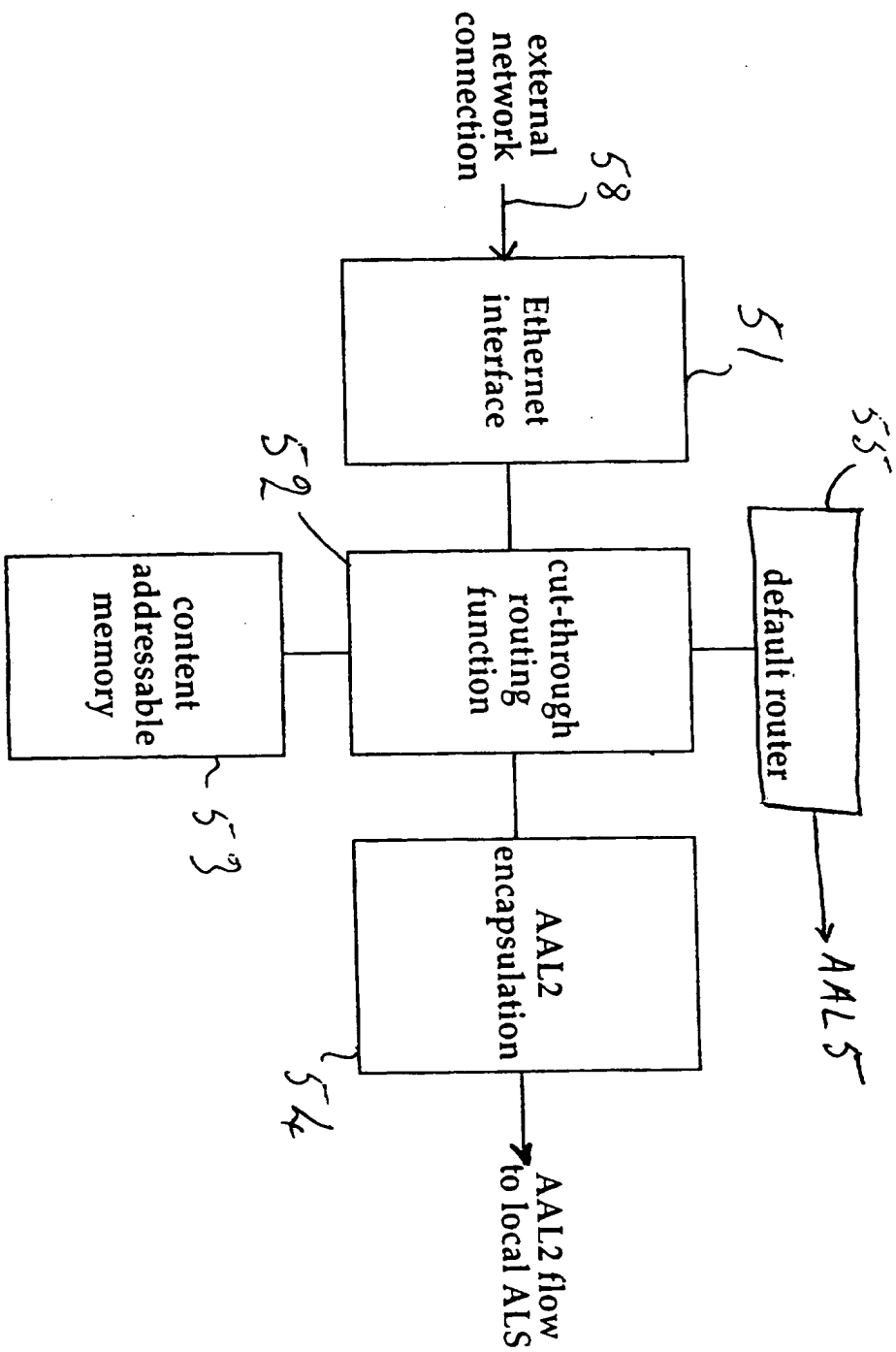


Figure 3

Figure 4



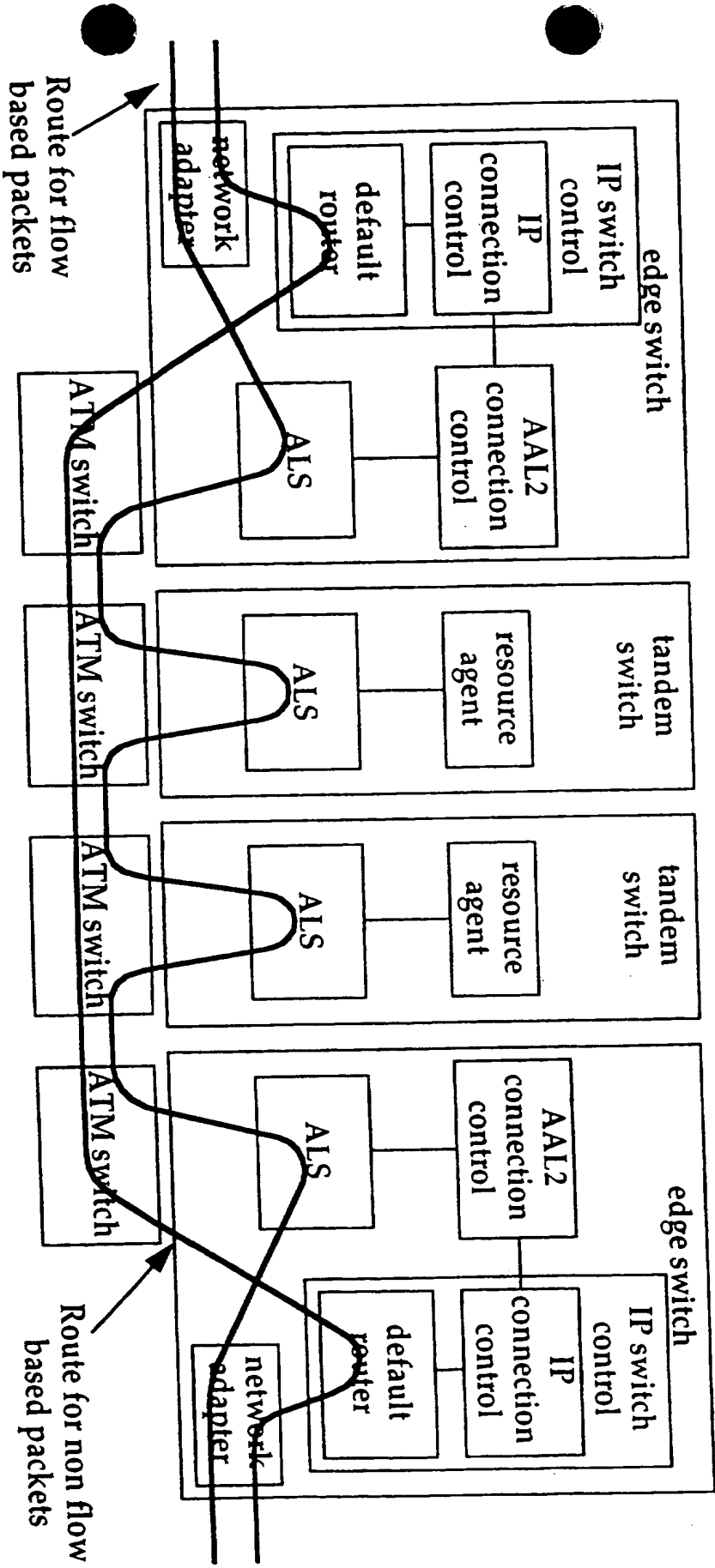
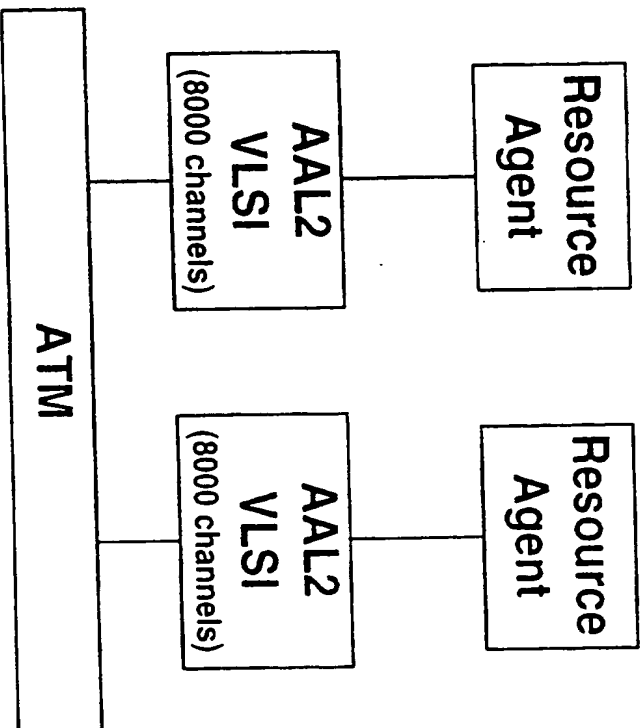


Figure 5

Figure 6



Adaptation Layer Switch Architecture

2320396

SCALEABLE DATA NETWORK ROUTER

This invention relates to communications networks and in particular to an arrangement and method for the routing of connectionless packet data in such networks.

5 BACKGROUND TO THE INVENTION

Data networks based on packet protocols, such as Internet Protocol (IP), use addressing information in the packet header to determine the destination of individual packets. Since each packet is treated independently of any other packet, the networks are connectionless.

10 Packets pass individually through the network between router nodes which determine to which router the packet should be sent next. The processing power required to interpret packet header information and determine the route to the next network node limits the capacity of the router.

15

In many instances, a communication between two terminal nodes in a communications network involves transmitting many IP packets in sequence. This is commonly referred to as a 'flow'. When a flow is identified by a router, a temporary link 'connection' can be set up
20 through the network to carry this flow. All packets with the same header information can then pass over this connection, avoiding the need to process the headers on an individual basis. This technique is known as IP switching.

25 With the introduction of broad band asynchronous transfer mode (ATM) networks, there is now a need for the transport of connectionless packet traffic over such networks, particularly with the introduction of IP voice services to provide Internet based telephony. However, Internet traffic is not readily adapted to real time communication over a broad band
30 network owing to its connectionless format, its multiple routing hops and its typical packet length which does not match the standard ATM cell.

SUMMARY OF THE INVENTION

An object of the invention is to minimise or to overcome this disadvantage.

5

A further object of the invention is to provide an improved method of carrying IP packet traffic over abroad band network.

10

The present invention relates to the use of IP switching techniques to implement a scaleable distributed router using ATM Adaptation Layer 2 (AAL2) as the bearer for IP packets identified as part of a flow.

15

According to one aspect of the invention there is provided a distributed router consisting of a plurality of edge switches interconnected by a plurality of tandem switches where flow based traffic is identified in the edge switch and a temporary connection is established to the destination edge switch where the packet is reconstituted into its original form for onward transmission.

20

According to another aspect of the invention there is provided an implementation of Internet Protocol (IP) real time services in which the end to end delay is fixed and low enough to achieve ITU delay standards for voice services.

25

According to a further aspect of the invention there is provided a method of routing traffic in a distributed router consisting of a plurality of edge switches interconnected by a plurality of tandem switches, the method including identifying flow based traffic in the edge switch and establishing a temporary connection into the destination edge switch whereby the packet is reconstituted into its original form for onward transmission.

30

According to another aspect of the invention, there is provided a method of routing packet data traffic in a broad band network comprising a plurality of edge switches interconnected by a plurality of tandem switches, the method including determining from each said packet whether that packet forms part of a flow of packets, where a said packet is determined to be part of a flow, segmenting that packet into minicells for transmission over a permanent virtual channel established across the network for that flow, and where a said packet is not determined to be

35

part of a flow, identifying a destination edge switch for that packet and transporting the packet to that edge switch via a default permanent virtual channel.

- 5 According to a further aspect of the invention there is provided a broad band network comprising a plurality of edge switches interconnected by a plurality of tandem switches and having means for routing packet data traffic, the network including means for determining from each said packet whether that packet forms part of a flow of packets, means for
10 segmenting a said packet determined to be part of a flow into minicells for transmission over a permanent virtual circuit established across the network for that flow, and means for identifying a destination edge switch for a said packet and transporting that packet to that edge switch via a default permanent virtual channel .where that packet is not determined
15 to be part of a flow.

A flow is defined as an established sequence of IP packets from a source to a destination comprising e.g. a file transfer or an Internet voice call.

20

In our arrangement, the ATM network functions as a distributed router whereby the IP packets are routed to the correct destination. Where a packet is identified as forming part of a flow, the destination will have already been established so that the packet payload can be carried in
25 minicells over the permanent virtual channel that has been provisioned for that flow.

- In a preferred embodiment, routing for a flow is determined from the network periphery by defining a succession of nodes and virtual
30 channels (VCs) to be used to reach the destination of that flow.

BRIEF DESCRIPTION OF THE DRAWINGS

An embodiment of the invention will now be described with respect to the following drawings in which:

35

Figure 1 is an overview of a data network arrangement according to an embodiment of the invention;

Figure 2 is a block diagram of edge and tandem switches on an example route through the data network of figure 1;

Figure 3 illustrates the general construction of an IP packet;

5

Figure 4 is an example realisation of an adaptation layer switch in the network of figure 1;

Figure 5 is a block schematic diagram of the network adapter shown in figure 2; and

10

Figure 6 shows example routes through the network of figure 1 for flow based and non flow based packets.

15 **DESCRIPTION OF PREFERRED EMBODIMENT**

Referring to figure 1, the broad band network comprises a plurality of interconnected nodes which nodes are of two types, edge switches 11 and tandem switches 12. Edge switches interconnect and interface with other elements of a data network and are themselves interconnected via

20

In the network of figure 1, the nodes are interconnected by permanent virtual channels (PVCs) 13 of a data network based on asynchronous transfer mode (ATM) technology. The ability of an ATM network to support PVCs between pairs of nodes on a network is well known. The PVCs are set up between each edge switch 11 and one or more tandem switches 12. Further PVCs are established between tandem switches 12 such that a highly or fully meshed core network is generated. Where each edge switch is connected to at least two tandem switches, then at least two possible routes will exist from any source edge switch to any other destination edge switch across the network.

25

30

35

In figure 1, for the purpose of illustration, traffic between any source edge switch passes to any destination edge switch through at most two tandem switches. It will be appreciated however that in a large practical network, traffic might pass through many tandem switches before reaching its destination edge switch.

Figure 2 shows the general construction and interconnection of the edge switches and tandem switches. Each edge switch 11 incorporates an IP network adaptor 21 providing an interface with an IP network employing e.g. an Ethernet protocol. The adaptor 21 is coupled to an IP switch control 22 and, via an adaptation layer switch (ALS) 23 to a connection control 24 and to an ATM switch 25. The connection control determines the establishment of a PVC for a flow. Each tandem switch 12 comprises a resource agent 26 coupled via an adaptation layer switch 23a to an ATM switch 25a.

Figure 3 shows in schematic form the construction of a typical IP data packet comprising a payload 30 and a header 31 containing supervisory information. The header comprises a number of information fields including a packet length indicator (LI), a protocol indicator (PI), a source identity (SI) and a destination identity (DI). Other header fields (not shown) may be provided to indicate e.g. the nature of the payload content, a sequence number and link layer information.

Figure 4 illustrates the network adapter function block of the edge switch of figures 1 and 2. In this example, traffic in the form of data packets arrives at an Ethernet interface 51 from the externally connected data network on an Ethernet connection 58. After removal of the Ethernet link layer information, the data packet passes to the cut-through routing function 52. The routing function 52 compares the packet header information with the entries from previous packets stored in a content addressable memory 53 to ascertain whether or not the incoming packet belongs to an already identified flow. If this is the case, the packet is segmented and encapsulated via an encapsulation function 54 in ATM adaptation layer 2 (AAL2) minicells for transmission over the connection which has been set up for that flow. The existence of a flow may be determined for example by the identification of a number of IP packets having the same source and destination identifiers. These identifiers from the first packet or packets of a suspected flow can be stored in the content addressable memory for comparison with corresponding information from subsequent packets.

If the packet is not recognised as belonging to a flow, the packet is passed to the default router 55 which processes the packet header to determine the distributed router output port (i.e. the edge switch) to which the packet should be sent. The packet is then segmented and
5 sent over an ATM PVC to the destination port, using the underlying ATM network in the normal way. ATM adaptation layer 5 (AAL5) is to be preferred for this purpose. . The respective routing of flow based and non flow based traffic through the network is illustrated in figure 5.

During processing of the packet header, the default router may
10 determine that this packet is (or is likely to be) the first part of a flow. If so, packet addressing information is passed to the IP connection control 41 (figure 4) via the default router. The IP connection control then processes the packet addressing information to determine its destination and requests the AAL2 connection control to establish a route for the
15 flow through the AAL2 adaptation layer switch (ALS) network. The function of the ALS is described in B-ISDN ATM Adaptation Layer Type 2 Specification I363.2. The route may be established using a technique called 'Wormhole Routing'. A detailed description of this routing technique is given in our co-pending United Kingdom patent application
20 No 9614138.7.

Although not shown in the drawings, it will be appreciated that sequence buffering may be provided at the network periphery to ensure the correct ordering of the first few packets of a flow. Once the flow is established, the use of the permanent virtual channel allocated to that flow ensures
25 that subsequent packets are maintained in sequence over the ATM network between the transmitting and receiving edge switches.

Wormhole Routing is a technique whereby the connection admission control (CAC) function is performed on the periphery of the network leading to a connection specification or 'worm' (figure 2) which defines
30 the succession of the nodes and the AAL2 VC's to be used to reach the endpoint. The CAC function is supported by resource agents built into the ALS components which monitor the usage of resources and report any impending exhaustion to the CAC functions at the periphery. This technique furnishes a guaranteed route through the ALS network,
35 choosing between multiple path options to take account of congestion

and path availability. Once this path is established, an entry is made in the content addressable memory 53 so that future packets belonging to the new flow will then be recognised and in consequence be correctly directed via the ALS network.

5

One possible implementation of an ALS node 23 (figure 2) is illustrated in figure 6, the AAL2 function is implemented as a VLSI component 231 serving typically 8000 channels. Multiple instances of this component are able to work over an ATM switch to achieve an ALS of $n \times 8000$ channels. In one particular embodiment 'n' can range from 1 to 60 so that an ALS node of about 500,000 channels is possible. The resource agent function 26 is implemented in a processor associated with each VLSI component to implement the 'wormhole routing' within as well as between nodes.

15

At the far end of the link, flow based traffic is extracted from the AAL2 minicell stream and reassembled into IP packets in the terminating network adapter. The link layer information is added and the packet exits the distributed router via the terminating Ethernet interface.

20

In our arrangement and method, routing intelligence is distributed around the edge switches; intelligence is added as edge switches are added. The distributed router described therefore scales to very large sizes, and could be implemented to global dimensions. For example a network of 2000 edge switches and 200 tandem switches has the potential to provide 500 million customers with 0.1 Erlangs of communication sessions with up to 5 flows in each session.

The value of AAL2 in this context is that its minicell structure eliminates the cell assembly delay normally associated with ATM. This is a well known problem at lower bit rates where the time required to fill an ATM cell before it can be launched may be unacceptably long. As an example, consider an IP flow comprising a voice service using G729 coding then G729 is an 8kb/s compression scheme which has a 10 byte frame and an encoding function which requires about 10 msec of processing. If this IP flow were to be packed into an AAL1 cell at four frames to a cell, then the cell assembly function would take 50 msec and a one cell receive buffer at the far end would add an additional 20 msec.

30

35

This 70 msec delay inherent in the cell assembly process is a large part of the ITU objective of a maximum 150 msec delay for terrestrially routed voice calls. By contrast, the use of AAL2 in the arrangement and method described above allows for example a 10 byte voice frame to be
5 launched at the end of compression processing, typically taking 20 msec, and the minicell itself can form the receive buffer delay of 5 msec. The delay of the ALS switches is fixed and substantially less than 1 msec, giving a total delay of about 25msec rather than the 70 msec that was previously required. Thus, the use of AAL2 in our arrangement and
10 method allows ATM to support compressed voice services with much lower latency and so allows IP voice services to achieve PSTN voice quality.

It will be understood that the above description of a preferred
15 embodiment is given by way of example only and that various modifications may be made by those skilled in the art without departing from the spirit and scope of the invention.

CLAIMS

1. A method of routing traffic in a distributed router arrangement
5 consisting of a plurality of edge switches interconnected by a plurality of tandem switches, the method including identifying flow based traffic in the edge switch and establishing a temporary connection into the destination edge switch whereby the packet is reconstituted into its original form for onward transmission.
10
2. A method of routing packet data traffic in a broad band network comprising a plurality of edge switches interconnected by a plurality of tandem switches, the method including determining from each said packet whether that packet forms part of a flow of packets, where a said
15 packet is determined to be part of a flow, segmenting that packet into minicells for transmission over a permanent virtual channel established across the network for that flow, and where a said packet is not determined to be part of a flow, identifying a destination edge switch for that packet and transporting the packet to that edge switch via a default
20 permanent virtual channel.
3. A method as claimed in claim 2, wherein the determination of whether a said packet is part of a flow is performed by comparison of
25 information stored in the header of that packet with corresponding information stored in a memory.
4. A method as claimed in claim 3, wherein a route for a new flow is established by defining on the periphery of the network a succession of nodes and virtual channels (VCs) to be used to reach the destination of
30 that flow.
5. A method as claimed in claim 2, 3 or 4, wherein the segmented packets constituting a flow are transported in ATM Adaptation Layer Two (AAL2) minicells.

6. A method as claimed in any one of claims 2 to 5, wherein said packets contain Internet voice traffic.

5 7. A method of routing traffic in a distributed router arrangement substantially as described herein with reference to and as shown in the accompanying drawings.

8. A distributed router arrangement consisting of a plurality of edge
10 switches interconnected by a plurality of tandem switches, wherein flow based traffic is identified in each said edge switch and a temporary connection is established to the corresponding destination edge switch where the packet is reconstituted into its original form for onward transmission.

15

9. A distributed router arrangement substantially as described herein with reference to and as shown in the accompanying drawings.

10. A broad band network comprising a plurality of edge switches
20 interconnected by a plurality of tandem switches and having means for routing packet data traffic, the network including means for determining from each said packet whether that packet forms part of a flow of packets, means for segmenting a said packet determined to be part of a flow into minicells for transmission over a permanent virtual circuit
25 established across the network for that flow, and means for identifying a destination edge switch for a said packet and transporting that packet to that edge switch via a default permanent virtual channel .where that packet is not determined to be part of a flow.

CLAIMS

1. A method of routing packet data traffic in a distributed router arrangement consisting of a plurality of edge switches interconnected by a plurality of tandem switches, the method including identifying flow based packet traffic in a first said edge switch, establishing a permanent virtual channel into a second edge switch, said second edge switch being a destination edge switch for that flow, whereby the packet is reconstituted into its original form for onward transmission, and transmitting said flow based packet traffic to the destination edge switch via the permanent virtual channel.
2. A method of routing packet data traffic in a broad band network comprising a plurality of edge switches interconnected by a plurality of tandem switches, the method including determining from each said packet whether that packet forms part of a flow of packets, where a said packet is determined to be part of a flow, segmenting that packet into minicells for transmission over a permanent virtual channel established across the network for that flow, and where a said packet is not determined to be part of a flow, identifying a destination edge switch for that packet and transporting the packet to that edge switch via a default permanent virtual channel.
3. A method as claimed in claim 2, wherein the determination of whether a said packet is part of a flow is performed by comparison of information stored in the header of that packet with corresponding information stored in a memory.
4. A method as claimed in claim 3, wherein a route for a new flow is established by defining on the periphery of the network a succession of nodes and virtual channels (VCs) to be used to reach the destination of that flow.

5. A method as claimed in claim 2, 3 or 4, wherein the segmented packets constituting a flow are transported in ATM Adaptation Layer Two (AAL2) minicells.

5 6. A method as claimed in any one of claims 2 to 5, wherein said packets contain Internet voice traffic.

7. A method of routing traffic in a distributed router arrangement substantially as described herein with reference to and as shown in the
10 accompanying drawings.

8. A distributed router arrangement consisting of a plurality of edge switches interconnected by a plurality of tandem switches, wherein flow based packet traffic is identified in a first said edge switch and a permanent
15 virtual channel is established into a second edge switch, said second edge switch being a destination edge switch for that flow, whereby to effect transmission of said flow based packet traffic to the destination edge switch via the permanent virtual channel.

9. A distributed router arrangement substantially as described herein
20 with reference to and as shown in the accompanying drawings.

10. A broad band network comprising a plurality of edge switches interconnected by a plurality of tandem switches and having means for routing packet data traffic, the network including means for determining from
25 each said packet whether that packet forms part of a flow of packets, means for segmenting a said packet determined to be part of a flow into minicells for transmission over a permanent virtual circuit established across the network for that flow, and means for identifying a destination edge switch for a said packet and transporting that packet to that edge switch via a default
30 permanent virtual channel where that packet is not determined to be part of a flow.



Application No: GB 9725483.3
Claims searched: All

Examiner: Al Strayton
Date of search: 25 February 1998

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.P): H4K: KTK

Int Cl (Ed.6): H04L

Other: ONLINE: WPI; INSPEC

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	GB 2 254 529 A (GPT)	
A	EP 0 570 630 A1 (ALCATEL)	
A	EP 0 552 385 A1 (IBM)	

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.